

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/563,166
Source: IFWP
Date Processed by STIC: 1/13/06

ENTERED



IFWP

RAW SEQUENCE LISTING

DATE: 01/13/2006

PATENT APPLICATION: US/10/563,166

TIME: 10:52:05

Input Set : E:\11711-001-999 (Sequence).txt

Output Set: N:\CRF4\01132006\J563166.raw

5 <110> APPLICANT: Hidai, Chiaki
 7 <120> TITLE OF INVENTION: Protein Capable of Deposition onto Extracellular Matrix
 9 <130> FILE REFERENCE: 11711-001-999(P03-0057PCT)
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/563,166
 C--> 12 <141> CURRENT FILING DATE: 2005-12-29
 14 <150> PRIOR APPLICATION NUMBER: JP2003-188598
 15 <151> PRIOR FILING DATE: 2003-06-30
 17 <160> NUMBER OF SEQ ID NOS: 26
 19 <170> SOFTWARE: PatentIn version 3.2
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 2303
 23 <212> TYPE: DNA
 24 <213> ORGANISM: Mus musculus
 27 <220> FEATURE:
 28 <221> NAME/KEY: CDS
 29 <222> LOCATION: (619)..(2061)
 31 <400> SEQUENCE: 1
 32 gaattccggtaaactgagga caaaggtaatcggatgttgcatttttgcatttc 60
 34 ttccccagtgg ccttgatatt taaaactgatt cctgccacca ggtccttggcc 120
 36 cctgcgtctc atatttctgc atgctgcttt gtttgtatat agtgcgctcc 180
 38 ctgcgtccccc tccagctctc gcttcattgt tctccaaatc aagaaggcccc 240
 40 cgccagcagcg tgagccgtatcactgctgg ccgcattcgcc tgcgtgcgcg 300
 42 ggggagccag gaacccaaagg agccgcgc tgcgcgtgt gcctctgcta 360
 44 agcccccagcc tctctcaagc gcacccaccc ccgcgcaccc cagctcaggc 420
 46 tgagggtgaa tcacccttcc tctaggccca ccactttt atcgcccttc ccaagatgg 480
 48 agaagcgtcg cgggagggaaa gacgtcctct tgatctctga cagggcgcccc 540
 50 tcctgcaggc gcgcctcgcc tactgtgcc tccgctacga ccccgacca gcccagggtca 600
 52 cgtccgttag aagggtatc atg aag cac ttg gta gca gcc tgg ctt ttg gtt 651
 53 Met Lys His Leu Val Ala Ala Trp Leu Leu Val
 54 1 5 10
 56 gga ctc agc ctc ggg gtg ccc cag ttc ggc aaa ggt gac att tgc aac 699
 57 Gly Leu Ser Leu Gly Val Pro Gln Phe Gly Lys Gly Asp Ile Cys Asn
 58 15 20 25
 60 ccg aac ccc tgt gaa aat ggt ggc atc tgt ctg tca gga ctg gct gat 747
 61 Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Ser Gly Leu Ala Asp
 62 30 35 40
 64 gat tcc ttt tcc tgt gag tgt cca gaa ggc ttc gca ggt ccg aac tgc 795
 65 Asp Ser Phe Ser Cys Glu Cys Pro Glu Gly Phe Ala Gly Pro Asn Cys
 66 45 50 55
 68 tct agt gtt gtg gag gtt gca tca gat gaa gaa aag cct act tca gca 843
 69 Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Lys Pro Thr Ser Ala
 70 60 65 70 75
 72 ggt ccc tgc atc cct aac cca tgc cat aac gga gga acc tgt gag ata 891

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/563,166

DATE: 01/13/2006
TIME: 10:52:05

Input Set : E:\11711-001-999 (Sequence).txt
Output Set: N:\CRF4\01132006\J563166.raw

73 Gly Pro Cys Ile Pro Asn Pro Cys His Asn Gly Gly Thr Cys Glu Ile			
74 80 85 90			
76 agc gaa gcc tat cga gga gac aca ttc ata ggc tat gtt tgt aaa tgt			939
77 Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr Val Cys Lys Cys			
78 95 100 105			
80 cct cgg gga ttt aat ggg att cac tgt cag cac aat ata aat gaa tgt			987
81 Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn Ile Asn Glu Cys			
82 110 115 120			
84 gaa gct gag cct tgc aga aat ggc gga ata tgt acc gac ctt gtt gct			1035
85 Glu Ala Glu Pro Cys Arg Asn Gly Gly Ile Cys Thr Asp Leu Val Ala			
86 125 130 135			
88 aac tac tct tgt gaa tgc cca gga gaa ttt atg gga cga aat tgt caa			1083
89 Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly Arg Asn Cys Gln			
90 140 145 150 155			
92 tat aaa tgc tct ggg cca ttg gga atc gaa ggt ggg atc ata tct aat			1131
93 Tyr Lys Cys Ser Gly Pro Leu Gly Ile Glu Gly Gly Ile Ile Ser Asn			
94 160 165 170			
96 cag caa atc aca gct tca tct act cac cga gct ctt ttt gga ctc cgg			1179
97 Gln Gln Ile Thr Ala Ser Ser Thr His Arg Ala Leu Phe Gly Leu Arg			
98 175 180 185			
100 aag tgg tat ccc tac tat gct cga ctt aat aag aag ggc ctt ata aat			1227
101 Lys Trp Tyr Pro Tyr Tyr Ala Arg Leu Asn Lys Lys Gly Leu Ile Asn			
102 190 195 200			
104 gcc tgg aca gct gct gaa aat gac aga tgg cca tgg att cag ata aat			1275
105 Ala Trp Thr Ala Ala Glu Asn Asp Arg Trp Pro Trp Ile Gln Ile Asn			
106 205 210 215			
108 ttg caa aga aaa atg aga gtc act ggt gtt att acc caa gga gca aaa			1323
109 Leu Gln Arg Lys Met Arg Val Thr Gly Val Ile Thr Gln Gly Ala Lys			
110 220 225 230 235			
112 agg att gga agc cca gag tac ata aaa tcc tac aaa att gcc tac agc			1371
113 Arg Ile Gly Ser Pro Glu Tyr Ile Lys Ser Tyr Lys Ile Ala Tyr Ser			
114 240 245 250			
116 aat gac ggg aag acc tgg gca atg tac aaa gta aaa ggc acc aat gaa			1419
117 Asn Asp Gly Lys Thr Trp Ala Met Tyr Lys Val Lys Gly Thr Asn Glu			
118 255 260 265			
120 gag atg gtc ttt cgt gga aat gtt gat aac aac aca cca tat gct aat			1467
121 Glu Met Val Phe Arg Gly Asn Val Asp Asn Asn Thr Pro Tyr Ala Asn			
122 270 275 280			
124 tct ttc aca ccc cca atc aaa gct cag tat gta aga ctc tac ccc caa			1515
125 Ser Phe Thr Pro Pro Ile Lys Ala Gln Tyr Val Arg Leu Tyr Pro Gln			
126 285 290 295			
128 att tgt cga agg cat tgt act tta aga atg gaa ctt ctt ggc tgt gag			1563
129 Ile Cys Arg Arg His Cys Thr Leu Arg Met Glu Leu Leu Gly Cys Glu			
130 300 305 310 315			
132 ctc tca ggc tgt tca gaa cct ttg ggg atg aaa tca ggg cat ata caa			1611
133 Leu Ser Gly Cys Ser Glu Pro Leu Gly Met Lys Ser Gly His Ile Gln			
134 320 325 330			
136 gac tac cag atc act gcc tcc agc gtc ttc aga aca ctc aac atg gac			1659
137 Asp Tyr Gln Ile Thr Ala Ser Ser Val Phe Arg Thr Leu Asn Met Asp			

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/563,166

DATE: 01/13/2006
TIME: 10:52:05

Input Set : E:\11711-001-999 (Sequence).txt
Output Set: N:\CRF4\01132006\J563166.raw

138	335	340	345	
140 atg ttt act tgg gaa cca agg aaa gcc agg ctg gac aag caa ggc aaa				1707
141 Met Phe Thr Trp Glu Pro Arg Lys Ala Arg Leu Asp Lys Gln Gly Lys				
142 350	355	360		
144 gta aat gcc tgg act tcc ggc cat aac gac cag tca caa tgg tta cag				1755
145 Val Asn Ala Trp Thr Ser Gly His Asn Asp Gln Ser Gln Trp Leu Gln				
146 365	370	375		
148 gtt gat ctt ctt gtc cct act aag gtg aca ggc atc att aca caa gga				1803
149 Val Asp Leu Leu Val Pro Thr Lys Val Thr Gly Ile Ile Thr Gln Gly				
150 380	385	390	395	
152 gct aaa gat ttt ggt cac gtg cag ttt gtt ggg tca tac aaa cta gct				1851
153 Ala Lys Asp Phe Gly His Val Gln Phe Val Gly Ser Tyr Lys Leu Ala				
154 400	405	410		
156 tac agc aat gat gga gaa cac tgg atg gtg cac cag gat gaa aaa cag				1899
157 Tyr Ser Asn Asp Gly Glu His Trp Met Val His Gln Asp Glu Lys Gln				
158 415	420	425		
160 agg aaa gac aag gtt ttt caa ggc aat ttt gac aat gac act cac agg				1947
161 Arg Lys Asp Lys Val Phe Gln Gly Asn Phe Asp Asn Asp Thr His Arg				
162 430	435	440		
164 aaa aat gtc atc gac cct ccc atc tat gca cga ttc ata aga atc ctt				1995
165 Lys Asn Val Ile Asp Pro Pro Ile Tyr Ala Arg Phe Ile Arg Ile Leu				
166 445	450	455		
168 cct tgg tcc tgg tat gga agg atc act ctg cgg tca gag ctg ctg ggc				2043
169 Pro Trp Ser Trp Tyr Gly Arg Ile Thr Leu Arg Ser Glu Leu Leu Gly				
170 460	465	470	475	
172 tgc gca gag gag gaa tga agtgcggggc cgcacatccc acaatgcttt				2091
173 Cys Ala Glu Glu				
174 480				
176 tctttatttt cctataagta tctccacgaa atgaactgtg tgaagctgat ggaaactgca				2151
178 ttgtttttt tcaaagtgtt caaattatgg taggctactg actgtcttt taggagttct				2211
180 aagcttgcc tttaataat ttaatttggt ttcccttgct caactcttctt atgtaataatc				2271
182 acactgtctg tgagttactc ttcttggtct ct				2303
185 <210> SEQ ID NO: 2				
186 <211> LENGTH: 480				
187 <212> TYPE: PRT				
188 <213> ORGANISM: Mus musculus				
190 <400> SEQUENCE: 2				
192 Met Lys His Leu Val Ala Ala Trp Leu Leu Val Gly Leu Ser Leu Gly				
193 1 5 10 15				
196 Val Pro Gln Phe Gly Lys Gly Asp Ile Cys Asn Pro Asn Pro Cys Glu				
197 20 25 30				
200 Asn Gly Gly Ile Cys Leu Ser Gly Leu Ala Asp Asp Ser Phe Ser Cys				
201 35 40 45				
204 Glu Cys Pro Glu Gly Phe Ala Gly Pro Asn Cys Ser Ser Val Val Glu				
205 50 55 60				
208 Val Ala Ser Asp Glu Glu Lys Pro Thr Ser Ala Gly Pro Cys Ile Pro				
209 65 70 75 80				
212 Asn Pro Cys His Asn Gly Gly Thr Cys Glu Ile Ser Glu Ala Tyr Arg				
213 85 90 95				

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/563,166

DATE: 01/13/2006
TIME: 10:52:05

Input Set : E:\11711-001-999 (Sequence).txt
Output Set: N:\CRF4\01132006\J563166.raw

```

216 Gly Asp Thr Phe Ile Gly Tyr Val Cys Lys Cys Pro Arg Gly Phe Asn
217          100           105           110
220 Gly Ile His Cys Gln His Asn Ile Asn Glu Cys Glu Ala Glu Pro Cys
221          115           120           125
224 Arg Asn Gly Gly Ile Cys Thr Asp Leu Val Ala Asn Tyr Ser Cys Glu
225          130           135           140
228 Cys Pro Gly Glu Phe Met Gly Arg Asn Cys Gln Tyr Lys Cys Ser Gly
229          145           150           155           160
232 Pro Leu Gly Ile Glu Gly Ile Ile Ser Asn Gln Gln Ile Thr Ala
233          165           170           175
236 Ser Ser Thr His Arg Ala Leu Phe Gly Leu Arg Lys Trp Tyr Pro Tyr
237          180           185           190
240 Tyr Ala Arg Leu Asn Lys Lys Gly Leu Ile Asn Ala Trp Thr Ala Ala
241          195           200           205
244 Glu Asn Asp Arg Trp Pro Trp Ile Gln Ile Asn Leu Gln Arg Lys Met
245          210           215           220
248 Arg Val Thr Gly Val Ile Thr Gln Gly Ala Lys Arg Ile Gly Ser Pro
249          225           230           235           240
252 Glu Tyr Ile Lys Ser Tyr Lys Ile Ala Tyr Ser Asn Asp Gly Lys Thr
253          245           250           255
256 Trp Ala Met Tyr Lys Val Lys Gly Thr Asn Glu Glu Met Val Phe Arg
257          260           265           270
260 Gly Asn Val Asp Asn Asn Thr Pro Tyr Ala Asn Ser Phe Thr Pro Pro
261          275           280           285
264 Ile Lys Ala Gln Tyr Val Arg Leu Tyr Pro Gln Ile Cys Arg Arg His
265          290           295           300
268 Cys Thr Leu Arg Met Glu Leu Leu Gly Cys Glu Leu Ser Gly Cys Ser
269          305           310           315           320
272 Glu Pro Leu Gly Met Lys Ser Gly His Ile Gln Asp Tyr Gln Ile Thr
273          325           330           335
276 Ala Ser Ser Val Phe Arg Thr Leu Asn Met Asp Met Phe Thr Trp Glu
277          340           345           350
280 Pro Arg Lys Ala Arg Leu Asp Lys Gln Gly Lys Val Asn Ala Trp Thr
281          355           360           365
284 Ser Gly His Asn Asp Gln Ser Gln Trp Leu Gln Val Asp Leu Leu Val
285          370           375           380
288 Pro Thr Lys Val Thr Gly Ile Ile Thr Gln Gly Ala Lys Asp Phe Gly
289          385           390           395           400
292 His Val Gln Phe Val Gly Ser Tyr Lys Leu Ala Tyr Ser Asn Asp Gly
293          405           410           415
296 Glu His Trp Met Val His Gln Asp Glu Lys Gln Arg Lys Asp Lys Val
297          420           425           430
300 Phe Gln Gly Asn Phe Asp Asn Asp Thr His Arg Lys Asn Val Ile Asp
301          435           440           445
304 Pro Pro Ile Tyr Ala Arg Phe Ile Arg Ile Leu Pro Trp Ser Trp Tyr
305          450           455           460
308 Gly Arg Ile Thr Leu Arg Ser Glu Leu Leu Gly Cys Ala Glu Glu Glu
309          465           470           475           480
312 <210> SEQ ID NO: 3

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/563,166

DATE: 01/13/2006
TIME: 10:52:05

Input Set : E:\11711-001-999 (Sequence).txt
Output Set: N:\CRF4\01132006\J563166.raw

```

313 <211> LENGTH: 393
314 <212> TYPE: DNA
315 <213> ORGANISM: Mus musculus
318 <220> FEATURE:
319 <221> NAME/KEY: CDS
320 <222> LOCATION: (1)..(393)
322 <400> SEQUENCE: 3
323 ata aat ttg caa aga aaa atg aga gtc act ggt gtt att acc caa gga      48
324 Ile Asn Leu Gln Arg Lys Met Arg Val Thr Gly Val Ile Thr Gln Gly
325 1           5           10          15
327 gca aaa agg att gga agc cca gag tac ata aaa tcc tac aaa att gcc      96
328 Ala Lys Arg Ile Gly Ser Pro Glu Tyr Ile Lys Ser Tyr Lys Ile Ala
329   20          25          30
331 tac agc aat gac ggg aag acc tgg gca atg tac aaa gta aaa ggc acc      144
332 Tyr Ser Asn Asp Gly Lys Thr Trp Ala Met Tyr Lys Val Lys Gly Thr
333   35          40          45
335 aat gaa gag atg gtc ttt cgt gga aat gtt gat aac aac aca cca tat     192
336 Asn Glu Glu Met Val Phe Arg Gly Asn Val Asp Asn Asn Thr Pro Tyr
337   50          55          60
339 gct aat tct ttc aca ccc cca atc aaa gct cag tat gta aga ctc tac     240
340 Ala Asn Ser Phe Thr Pro Pro Ile Lys Ala Gln Tyr Val Arg Leu Tyr
341   65          70          75          80
343 ccc caa att tgt cga agg cat tgt act tta aga atg gaa ctt ctt ggc     288
344 Pro Gln Ile Cys Arg Arg His Cys Thr Leu Arg Met Glu Leu Leu Gly
345   85          90          95
347 tgt gag ctc tca ggc tgt tca gaa cct ttg ggg atg aaa tca ggg cat     336
348 Cys Glu Leu Ser Gly Cys Ser Glu Pro Leu Gly Met Lys Ser Gly His
349   100         105         110
351 ata caa gac tac cag atc act gcc tcc agc gtc ttc aga aca ctc aac     384
352 Ile Gln Asp Tyr Gln Ile Thr Ala Ser Ser Val Phe Arg Thr Leu Asn
353   115         120         125
355 atg gac atg
356 Met Asp Met
357   130
360 <210> SEQ ID NO: 4
361 <211> LENGTH: 131
362 <212> TYPE: PRT
363 <213> ORGANISM: Mus musculus
365 <400> SEQUENCE: 4
367 Ile Asn Leu Gln Arg Lys Met Arg Val Thr Gly Val Ile Thr Gln Gly
368 1           5           10          15
371 Ala Lys Arg Ile Gly Ser Pro Glu Tyr Ile Lys Ser Tyr Lys Ile Ala
372   20          25          30
375 Tyr Ser Asn Asp Gly Lys Thr Trp Ala Met Tyr Lys Val Lys Gly Thr
376   35          40          45
379 Asn Glu Glu Met Val Phe Arg Gly Asn Val Asp Asn Asn Thr Pro Tyr
380   50          55          60
383 Ala Asn Ser Phe Thr Pro Pro Ile Lys Ala Gln Tyr Val Arg Leu Tyr
384   65          70          75          80

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/563,166

DATE: 01/13/2006

TIME: 10:52:06

Input Set : E:\11711-001-999 (Sequence).txt

Output Set: N:\CRF4\01132006\J563166.raw

L:11 M:270 C: Current Application Number differs, Replaced Application Number

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date